Lecture

INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: SY BSc

Subject : Statistical and Risk Modelling 1
Subject Code: PUSASQF

Chapter: Unit 3 Chp 1

Chapter Name: Graduation 1

Today's Agenda

1. Graduation
2. Need for graduation
3. Origins of graduation and smoothing
4. Assumptions
5. Comparison with other experiences
6. Procedure to test comparison
7. Desirable features and issues of graduation
8. Testing for smoothness
    1. The test

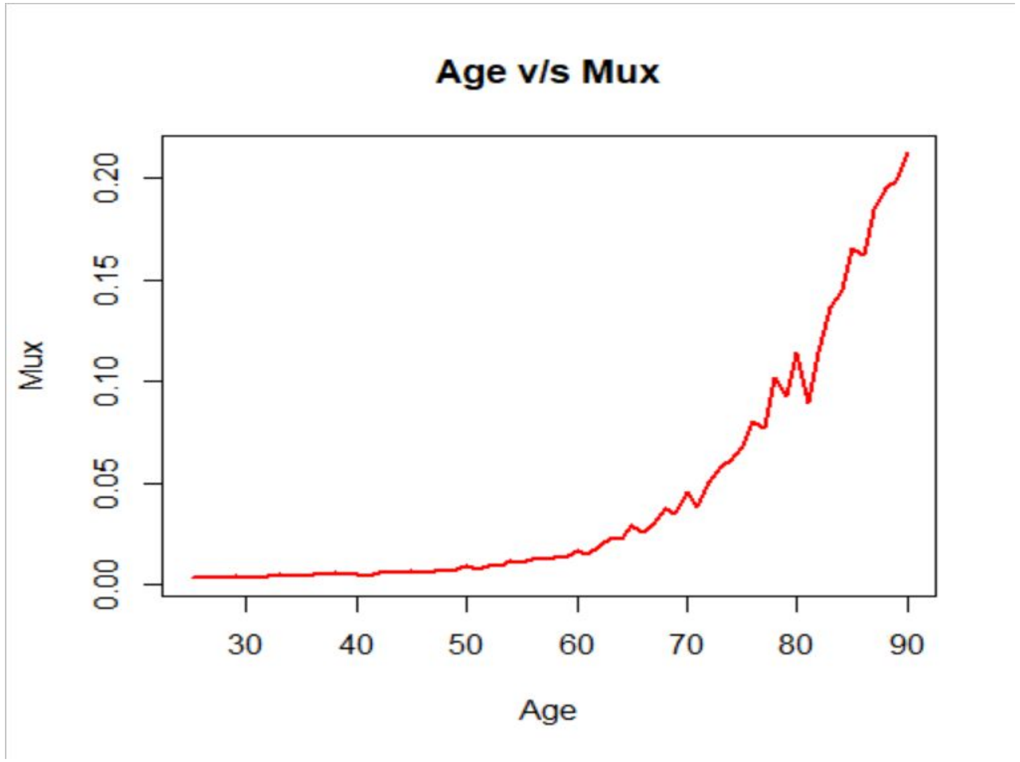
Continued …

9. Statistical test for testing graduation
    1. Chi-Squared test
    2. Limitations of Chi-Squared test
    3. Deviations and standardized deviations
    4. Standardized deviations test
    5. Signs test
    6. Cumulative Deviations test
    7. Grouping of Signs test
    8. Serials Correlations test

10. Methods for Graduation

Introduction

We’ve previously learnt about a few models and the way crude mortality rates are modelled. Now we’ll try to refine them for our use.



Crude Rates against Age

1 Graduation

Graduation is a technique used to get a smooth and justifiable mortality rates. (which is also called as Graduated Mortality Rate).

The crude mortality rates derived from a mortality investigation will not be the final rates that are published for use in actuarial calculations. The rates will have to pass through a further process called graduation. Graduation results in a 'smoothing' of the crude rates.

The aims of graduation are: –

- To produce a smooth set of rates that are suitable for a particular purpose.
- To remove random sampling error.
- To use information available from adjacent ages

2 Need for graduation

Let us understand why do we need to perform the process of graduation.

Theoretical argument:

i. We believe that mortality varies smoothly with age.
ii. Therefore the crude estimate at any age carries information about mortality at adjacent ages. i.e. a crude estimate of μ_x for any age x also carries information about the values of μ_{x-1} and μ_{x+1}.
iii. By smoothing, we can make use of the data at adjacent ages to improve the estimate at each age. Thus, it reduces sampling errors

2 Need for graduation

While theoretical arguments provide a good reason to graduate rates, the real reason why all companies perform graduation is not theoretical but practical reasons.

Practical argument:

- Mortality data is that which we will use in the life tables to compute financial quantities, such as premiums for life insurance contracts.
- It is very desirable that such quantities progress smoothly with age, since irregularities (jumps or other anomalies) are hard to justify in practice.
- Before applying the mortality rates to any financial problem, we need to consider their suitability. Thus we need to make comparisons with standard tables and other experiences. It is often the case that a mortality experience must be adjusted in some way before use, in which case there is little point in maintaining the roughness of the crude estimates.

3 Where it all began!!

Actuaries have had a very long association with smoothing. The Gompertz model (Gompertz, 1825) could be considered an early smoothing method, albeit a very simple one.

Makeham's extension (Makeham, 1860) of the Gompertz model improved the fit to mortality tables, but neither model was sufficiently flexible to be applicable outside a limited age range, say 40 to 90.

There were many other efforts, all grouped under the general heading of mathematical formulae, of which Perks (1932) is perhaps the best known. The basic idea is that adding a parameter will improve the fit, so we have the Gompertz model with two parameters, Makeham's with three and the two Perks formulae with four and five parameters.

English Life Tables No. 11 and No. 12 were graduated using a mathematical formula with seven parameters.

4 Assumptions

Consider that we have data for all ages in our investigation.

Using the Poisson or multiple-state model, for $x = x_1, x_2, \ldots\ldots x_m$ we have:

Number of deaths at age x nearest birthday = D_x

Central exposed to risk at age x nearest birthday = E_x^c

Crude estimate of the force of mortality at exact age = $\widehat{\mu_x}$

Approximate distribution that we use is

$$D_x \sim Normal(E_x^c \mu_x, E_x^c \mu_x)$$

Graduated rates are denoted as $\dot{\mu}_x$

5 Comparison with other experiences

- In mortality investigations and insurance studies, we often need to keep a check on the consistency of our data with other known experiences. It is important to see whether our recent experience is consistent with past experiences or with published life tables.

- By consistency, we cover two concepts: the shape of the mortality curve over the range of ages and the level of mortality rates.

Comparison with Standard tables.

- Published life tables based on large amounts of data are called standard tables. These mortality tables use a large number of factors to predict the likelihood of death in an individual. Mortality tables are used heavily by insurance companies.

6 Procedure to test comparison

- Notation: The superscript 's' will denote a quantity from a published standard table, *eg* μ_x^s
- Our hypothesis is that the mortality rates being tested are consistent with those from the standard table. We will reject this null hypothesis if we find evidence that the rates being tested are significantly different from those in the standard table.
- We can derive tests of this hypothesis using the distributional assumptions made earlier. Under the hypothesis we have,

$$\boldsymbol{D_x \sim N(E_x^c \mu_x^s, E_x^c \mu_x^s)} \quad \text{approximately}$$

- Thus we can find test statistics making comparisons of actual deaths d_x in our data and the expected deaths given by this distributional assumption There are suitable tests that can help us to work further

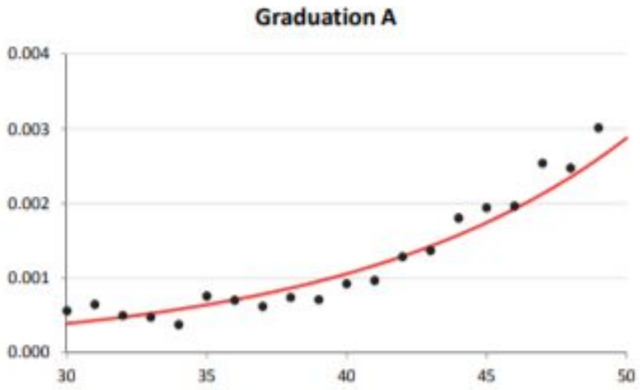
7 Desirable features and issues of Graduation

Three desirable features of a graduation are: –

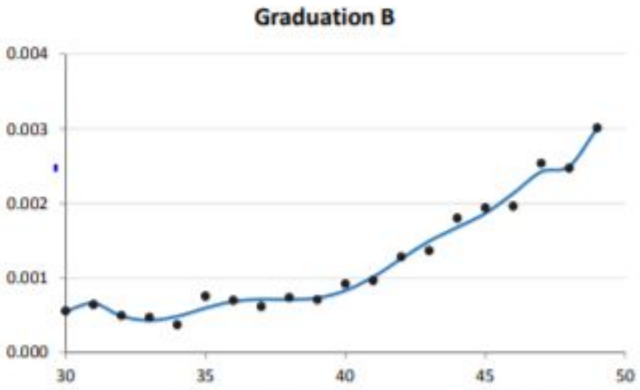
i. Smoothness
ii. Adherence to data
iii. Suitability for the purpose to hand

- We need to make a balance between smoothing and adherence to data. At one extreme, we could easily smooth the crude estimates by ignoring the data altogether; we want to avoid such extremes since we want the graduation to be representative of the experience.
- If the graduation process results in rates that are smooth but show little adherence to the data, then we say that the rates are over-graduated.
- If insufficient smoothing has been carried out then it is called under-graduation.

7 Desirable features and issues of Graduation



Graduation A - overgraduated



Graduation B - undergraduated

- Smoothness and adherence to data are usually conflicting requirements. Perfect smoothness (extreme example: a straight line) pays little or no attention to the data, while perfect adherence to the data means no smoothing at all.
- **Over graduation** – If the graduation process results in rates that are smooth but show little adherence to the data, then we say that the rates are overgraduated. The graph in Graduation A (see above) is very smooth, but it tends to overestimate the crude rates at the younger ages and underestimates them at the older ages.
- **Under graduation** – Graduation where insufficient smoothing has been carried out. This will tend to produce a curve of inadequate smoothness, but better adherence to data. The graph in Graduation B (which uses the same data as Graduation A) adheres very closely to the crude rates, but it twists and turns erratically

8 Testing for Smoothness

- The achievement of smoothness is one of the main purposes of graduation. The ‘art’ of graduation lies in finding a satisfactory compromise between smoothness and adherence to data.

Smooth Graduation

- Mathematical smoothness is usually defined in terms of differentiability, but this is of little use in graduation work because many functions that misbehave wildly between integer ages are nevertheless differentiable many times.
- The test for smoothness will be used as a check for undergraduation.

8.1 The test

To test smoothness we will find the third-order differences of the graduated rates.
The first difference $\Delta\dot{\mu}_x = \dot{\mu}_{x+1} - \dot{\mu}_x$
The second difference $\Delta^2\dot{\mu}_x = \Delta\dot{\mu}_{x+1} - \Delta\dot{\mu}_x$
The third difference $\Delta^3\dot{\mu}_x = \Delta^2\dot{\mu}_{x+1} - \Delta^2\dot{\mu}_x$

The criterion for the graduated rates to qualify the smoothness test is that the third differences should:

- be small in magnitude compared with the quantities themselves; and
- progress regularly.

Question 1

Compare the smoothness of the rates in Graduations A and B over the age range 30 to 35:
Looking at the graph in the next slide, also comment about undergraduation and overgraduation of rates

Graduation A

x	$\mathring{\mu}_x$	$\Delta \mathring{\mu}_x$	$\Delta^2 \mathring{\mu}_x$	$\Delta^3 \mathring{\mu}_x$
30	0.000388	0.000041	0.000004	0.000001
31	0.000429	0.000045	0.000005	0.000000
32	0.000474	0.000050	0.000005	0.000001
33	0.000524	0.000055	0.000006	
34	0.000579	0.000061		
35	0.000640			

Graduation B

x	$\mathring{\mu}_x$	$\Delta \mathring{\mu}_x$	$\Delta^2 \mathring{\mu}_x$	$\Delta^3 \mathring{\mu}_x$
30	0.000555	0.000103	–0.000273	0.000387
31	0.000658	–0.000170	0.000114	–0.000004
32	0.000488	–0.000056	0.000110	–0.000054
33	0.000432	0.000054	0.000056	
34	0.000486	0.000110		
35	0.000596			

Solution

For A, the third differences are very small, which indicates that the graduated rates are very smooth.
The third differences are larger for Graduation B than for Graduation A (especially when $x = 30$) and they progress in a less regular manner. This indicates that Graduation B is not as smooth as Graduation A.

Graduation A

We are going to use Graduation A & Graduation B for a few of the tests to follow. These charts represent their data.

Graduation A assumed that $\ln(e^{\mu_x} - 1)$ could be modelled as $\alpha + \beta x$ (2 parameters), which was fitted using the method of least squares.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c\,\mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c\mathring{\mu}_x}{\sqrt{E_x^c\mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	-0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	-2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	-0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	-1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	-1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	-2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	-1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	-1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	-0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	-0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Graduation B

We are going to use Graduation A & Graduation B for a few of the tests to follow. These charts represent their data.

Graduation B assumed that $\ln(e^{\mu_x} - 1)$ could be modelled as a polynomial of degree 10 (11 parameters), which was fitted using the method of least squares.

Graduation B

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000555	38.85	0.02	0.00
31	66,672	43	0.000645	0.000658	43.87	−0.13	0.02
32	68,375	34	0.000497	0.000488	33.37	0.11	0.01
33	65,420	31	0.000474	0.000432	28.26	0.52	0.27
34	61,779	23	0.000372	0.000486	30.02	−1.28	1.64
35	66,091	50	0.000757	0.000596	39.39	1.69	2.86
36	68,514	48	0.000701	0.000685	46.93	0.16	0.02
37	69,560	43	0.000618	0.000713	49.60	−0.94	0.88
38	65,000	48	0.000738	0.000709	46.09	0.28	0.08
39	66,279	47	0.000709	0.000733	48.58	−0.23	0.05
40	67,300	62	0.000921	0.000831	55.93	0.81	0.66
41	65,368	63	0.000964	0.001015	66.35	−0.41	0.17
42	65,391	84	0.001285	0.001259	82.33	0.18	0.03
43	62,917	86	0.001367	0.001494	94.00	−0.82	0.68
44	66,537	120	0.001804	0.001679	111.72	0.78	0.61
45	62,302	121	0.001942	0.001866	116.26	0.44	0.19
46	62,145	122	0.001963	0.002134	132.62	−0.92	0.85
47	63,856	162	0.002537	0.002423	154.72	0.59	0.34
48	61,097	151	0.002471	0.002498	152.62	−0.13	0.02
49	61,110	184	0.003011	0.003008	183.82	0.01	0.00
Total		1,561			1,555.31		9.39

9 Statistical tests for testing graduation

Statistical test process

Hypotheses - The null hypothesis (denoted by H_0) corresponds to a neutral conclusion. In graduation tests, the null hypothesis will correspond to a statement that some aspect of a proposed graduation is 'OK'. In graduation tests, the alternative hypothesis (H_1) will correspond to a statement that some aspect of a proposed graduation is 'no good'.

We then perform the required statistical test and find the probability value – (p value).

P – Value - A significance level must be selected at the beginning of the test. The significance level usually used is 5%, which means that if H_0 were true, the value of the test statistic would only be this extreme by chance 1 time in 20. A statistical test may be one-tailed or two-tailed, depending on the nature of the test and the feature we are interested in.

9.1 Chi-Squared Test

Chi-squared test is a test for the overall fitness of fit.
A chi-squared test can be used to assess whether the observed numbers of individuals in a specified categories are consistent with a model that predicts the expected numbers in each category.

Test Statistic - The chi-squared test statistic is: $\sum \frac{(O-E)^2}{E}$
where,
- O is the observed number in a particular category
- E is the corresponding expected number predicted by the assumed probabilities
- the sum is over all possible categories .

This statistic has a chi-squared distribution (approximately), which is tabulated in the statistics section of the Actuarial Tables.

Calculate Degrees of freedom as = No. of categories – number of parameters

Question 2

Perform the Chi Square Goodness of Fit test on the below data

Age	Crude Rates	Central Exposed to Risk	Graduated Rates
41	0.02	100	0.022
42	0.03	200	0.0305
43	0.033	150	0.034
44	0.067	120	0.068
45	0.07	100	0.075
46	0.075	100	0.08

One – parameter estimation has been performed to calculate the graduated rates.

Do you see an issue with your conclusion in the above test? If yes, what?

9.2 Limitations of Chi-Squared Test

The χ^2 test will fail to detect several defects that could be of considerable financial importance. (These comments apply particularly when we are testing a graduation, and for ease of exposition we will write as if that were the case.)

i. There could be a few large deviations offset by a lot of very small deviations. In other words, the χ^2 test could be satisfied although the data do not satisfy the distributional assumptions that underlie it. This is, in essence, because the χ^2 statistic summarises a lot of information in a single figure.
ii. The graduation might be biased above or below the data by a small amount. The χ^2 statistic can often fail to detect consistent bias if it is small, but we should still wish to avoid it.
iii. Even if the graduation is not biased as a whole, there could be significant groups of consecutive ages (called runs or clumps) over which it is biased up or down. This is still to be avoided.
iv. It should be noted that because the χ^2 test is based on squared deviations, it tells us nothing about the direction of any bias or the nature of any lack of adherence to data of a graduation, even if the bias is large or the lack of adherence manifest. To ascertain this there is no substitute for an inspection of the experience.

Question

Apply the chi-squared test to Graduation A.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \frac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	−0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	−2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	−0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	−1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	−1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	−2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	−1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	−1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	−0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	−0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Solution

From the table of values for Graduation A, we see that:

$$\sum(Z_X^2) = 43.17$$

In this example, it is not difficult to work out how many degrees of freedom to use. There are 20 ages. We have not constrained the totals. The graduated rates have been calculated by estimating 2 parameters. So, the number of degrees of freedom is 20 – 2 = 18 .

From the Tables, the upper 5% point for the χ_{18}^2 distribution is 28.87. The observed value of the test statistic exceeds this, so we reject the null hypothesis. (In fact, the test statistic also exceeds 42.31, the upper 0.1% point.)

So, we conclude that the mortality experience does not conform to a formula of the type assumed in the graduation.

9.3 Deviations and standardized deviations

When we are testing the adherence to data of a graduation, we assume the normal distribution is used as an approximation to the Poisson distributions.

$$\boldsymbol{D_x \sim N(E_x^c \dot{\mu}_x, E_x^c \dot{\mu}_x)}$$

- Deviations and standardized deviations:

  The deviation at age x is defined to be: **Actual deaths – Expected deaths**

  i.e. $\boldsymbol{D_x - E_x^c \dot{\mu}_x}$

  and the standardized deviation, denoted z_x is:

$$\boldsymbol{Z_x = \frac{D_x - E_x^c \dot{\mu}_x}{\sqrt{E_x^c \dot{\mu}_x}}}$$

  In case of comparison with standard table replace $\dot{\mu}_x$ by μ_x^s , above.

9.4 Standardized deviations test

The test looks at the distribution of the values of the standardized deviations.

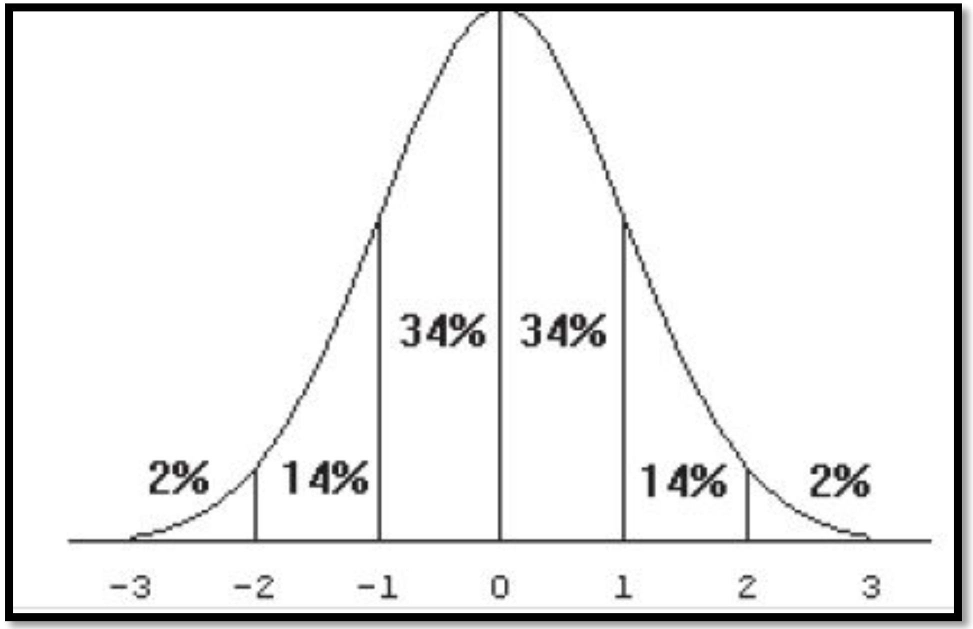
- **Step 1** – Calculate the standardized deviations z_x for each age .
- **Step 2** - Divide the real (number) line into any convenient intervals. Plot or count the number of standardized deviations falling into each of the ranges.
- **Step 3** - We can then compare: the observed number of the z_x that fall in each interval; and the expected number of the z_x that should fall in each interval, under the hypothesis.

The hypothesis states that the z_x values are realizations of a standard normal random variable. If the standardized deviations do not appear to conform to a standard normal distribution, this indicates that mortality rates do not conform to the model with the rates assumed in the graduation.

9.4 Standardized deviations test

In this example, the expected numbers are:

Interval		(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	
Expected number	0	0.02m	0.14m	0.34m	0.34m	0.14m	0.02m	0



9.4 Standardized deviations test

- **Step 4** -To formalise the comparison, we can form a χ^2 statistic (nothing to do with the use of the χ^2 test mentioned previously):

$$\chi^2_{cal} = \sum_{all\ intervals} \frac{(Actual - Expected)^2}{Expected}$$

which here should have a χ^2 distribution with 7 degrees of freedom (since we have used 8 intervals).

Note how this differs from the way we previously applied the χ^2 test, which was to test whether the observed numbers of deaths were consistent with a given set of graduated rates. Here, we are testing whether the observed pattern of the individual standardised deviations (ie the numbers falling in each interval) is consistent with a standard normal distribution.

- If the number of age groups is small, we should use a smaller number of intervals, ensuring that the expected number of standardised deviations in each interval is not less than five (recommended) but should not be less than one (as a rule of thumb), and we then reduce the number of degrees of freedom in the χ^2 test appropriately.

Question

Analyse the distribution of the standardised deviations for Graduation A.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \frac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	−0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	−2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	−0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	−1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	−1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	−2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	−1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	−1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	−0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	−0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Solution

The observed and expected numbers in each range are shown in the table below.

Interval	$(-\infty,-3)$	$(-3,-2)$	$(-2,-1)$	$(-1,0)$	$(0,1)$	$(1,2)$	$(2,3)$	$(3,\infty)$
Observed	0	2	4	4	3	4	3	0
Expected	0.0	0.4	2.8	6.8	6.8	2.8	0.4	0

There are only 7 values in the range (-2/3,2/3) . So, there appear to be too few values in the centre of the distribution and too many in the tails. This might indicate overgraduation (an inappropriate graduation formula) or the presence of duplicates.

The distribution of the standardised deviations is fairly symmetrical, with 10 positive and 10 negative values. So there is no evidence of bias in the graduated rates.

9.5 Signs test

The signs test is a simple test for overall bias. In other words, this test checks whether the graduated rates are too high or too low.

- **Step 1** - Count how many of the graduated rates lie above/below the crude rates. We will do this by looking at the signs of the individual standardized deviations.
- **Step 2** - Calculate the probability value for the test by finding the probability of obtaining a split of positive/negative values as extreme as observed.

Define the test statistic: P = Number of z_x that are positive
Under the hypothesis, P ~ Binomial(m,½) . So the probability function of P is:

$$\boldsymbol{P(P = x) = \binom{m}{x}\left(\frac{1}{2}\right)^{m}, x = 0, 1, \ldots, m}$$

An excess of either negative or positive deviations is a defect, so we apply a two-tailed test.

Question

State the conclusion that can be drawn from an examination of the signs of the deviations for Graduation B.

Graduation B

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000555	38.85	0.02	0.00
31	66,672	43	0.000645	0.000658	43.87	−0.13	0.02
32	68,375	34	0.000497	0.000488	33.37	0.11	0.01
33	65,420	31	0.000474	0.000432	28.26	0.52	0.27
34	61,779	23	0.000372	0.000486	30.02	−1.28	1.64
35	66,091	50	0.000757	0.000596	39.39	1.69	2.86
36	68,514	48	0.000701	0.000685	46.93	0.16	0.02
37	69,560	43	0.000618	0.000713	49.60	−0.94	0.88
38	65,000	48	0.000738	0.000709	46.09	0.28	0.08
39	66,279	47	0.000709	0.000733	48.58	−0.23	0.05
40	67,300	62	0.000921	0.000831	55.93	0.81	0.66
41	65,368	63	0.000964	0.001015	66.35	−0.41	0.17
42	65,391	84	0.001285	0.001259	82.33	0.18	0.03
43	62,917	86	0.001367	0.001494	94.00	−0.82	0.68
44	66,537	120	0.001804	0.001679	111.72	0.78	0.61
45	62,302	121	0.001942	0.001866	116.26	0.44	0.19
46	62,145	122	0.001963	0.002134	132.62	−0.92	0.85
47	63,856	162	0.002537	0.002423	154.72	0.59	0.34
48	61,097	151	0.002471	0.002498	152.62	−0.13	0.02
49	61,110	184	0.003011	0.003008	183.82	0.01	0.00
Total		1,561			1,555.31		9.39

Solution

There are 12 positive and 8 negative values for Graduation B. Here we have more positive deviations than expected. So the p-value is:

$$2P(p \geq 12) = 2[1 - P(P \leq 11)] = 2[1 - 0.7483] = 0.5034$$

The value of 0.7483 comes from page 188 of the Tables.
Since this is (much) greater than 5%, there is very little evidence of bias in the graduated rates.

9.6 Cumulative Deviations test

The cumulative deviations test detects overall bias or long runs of deviations of the same sign.
Applying normal approximation to the deviation and standardizing gives us :

$$\frac{\left(\sum_{\forall x}(\boldsymbol{D_x} - \boldsymbol{E_x^c}\dot{\boldsymbol{\mu}}_{\boldsymbol{x}})\right)}{\sqrt{\sum_{\forall x}\boldsymbol{E_x^c}\dot{\boldsymbol{\mu}}_{\boldsymbol{x}}}} \sim \boldsymbol{N}(\boldsymbol{0},\boldsymbol{1})$$

- Step 1 - Calculate ∑dx (the total observed deaths) and $E\sum_x^c \dot{\mu}_X$ (the total expected deaths), where the sum is over the selected age range.
- Step 2 - Calculate the test statistic $\frac{\left(\sum \boldsymbol{d_x} - \sum(\boldsymbol{E_X^c}\dot{\boldsymbol{\mu}}_{\boldsymbol{x}})\right)}{\sqrt{\sum(\boldsymbol{E_X^c}\dot{\boldsymbol{\mu}}_{\boldsymbol{x}})}}$ and use this to determine the p-value using the tables for the standard normal distribution.

Question

State the conclusion that can be drawn from applying the cumulative deviations test to the whole age range of Graduation A.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	−0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	−2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	−0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	−1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	−1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	−2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	−1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	−1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	−0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	−0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Solution

From the table given on Graduation A, the value of the test statistic is:

$$\frac{\{\sum d_X - \sum(E_X^c \dot{\mu}_X)]}{\sqrt{\sum(E_X^c \dot{\mu}_X)}} = \frac{1561 - 1516.50}{\sqrt{1516.50}} = 1.143$$

This is a two-tailed test, so we compare the value of the test statistic with the upper and lower 2.5% points of N(0,1) , ie ± 1.96 . As $-1.96 < 1.143 < 1.96$, there is insufficient evidence to reject the null hypothesis.

So, the cumulative deviations test does not provide evidence that the graduated rates are biased.

9.7 Grouping of Signs test

This test is used mainly to test for over graduation. Lets look at how the test is done…

- **Step 1** - Determine the sign of the deviation at each age.
- **Step 2** - Count the number of groups of positive signs (= G).
- **Step 3** - The hypothesis is that the given n_1 positive deviations and n_2 negative deviations are in random order. Since every pair of positive groups must be separated by a negative group, the numbers of positive and negative groups will be small or large alike, so a one-tailed test is appropriate. We should find the smallest k such that:

$$\sum_{t=1}^{k} \frac{\binom{n_1 - 1}{t - 1}\binom{n_2 + 1}{t}}{\binom{m}{n_1}} \geq 0.05$$

and say that the test has been failed (at the 5% level) if G < k .
If there are too few runs, this indicates that the rates are over graduated.

Alternatively, we could look up the critical value of the test on page 189 of the Tables. If the number of groups of positive deviations is less than or equal to the critical value given in the Tables, we reject the null hypothesis.

Question

Test Graduation A for overgraduation using the grouping of signs test.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	-0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	-2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	-0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	-1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	-1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	-2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	-1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	-1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	-0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	-0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Solution

Here there are 20 age groups, with 10 positive and 10 negative deviations. From page 189 of the Tables, we see that the critical value is 3 when n_1 =10 and n_2 =10 . Looking at the column of z_X values on page 12, we see that there are 3 runs of positive deviations. So we reject the null hypothesis at the 5% significance level and conclude that there is evidence of grouping of deviations of the same sign.

Alternatively, we could carry out the test using a normal approximation. The expected number of positive runs is: $\frac{10(10+1)}{10+10} = 5.5$ and the variance is: $\frac{(10X10)^2}{(10+10)^3} = 1.25$

The p-value is P(G≤ 3) . Using N(5.5,1.25) as an approximation to the distribution of G and incorporating a continuity correction:

$$P(G \leq 3) \cong P\left(N(0,1) \leq \frac{3.5 - 5.5}{\sqrt{1.25}}\right) = \Phi(-1.78885) = 1 - \Phi(1.78885) = 3.7\%$$

Since the p-value is less than 5%, we reject the null hypothesis at the 5% level and conclude that there is evidence of grouping of deviations of the same sign.

9.8 Serial Correlations test

We have this test to test for clumping of deviations of the same sign. If clumping is present, then the graduation has the wrong shape.

- **Step 1** - Calculate the standardized deviations z_X for each age or age group.
- **Step 2** - Calculate the serial correlation coefficients using the formula:

$$\boldsymbol{r_j = \frac{\frac{1}{m-j}\sum_{i=1}^{m-j}(z_i - \bar{z})(Z_{i+j} - \bar{z})}{\frac{1}{m}\sum_{i=1}^{m}(z_i - \bar{z})^2}}$$

where $\bar{\boldsymbol{z}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{z_i}$ is the overall average of z_X for the m ages

We can use the fact that r_j must take values in the range -1 ≤ r_j ≤ 1 to check the calculations for reasonableness, j is the lag.

- **Step 3** - Multiply by $\sqrt{m}$ to obtain the value of the test statistic and compare this with the percentage points of the standard normal distribution.

  If the test statistic is 'too positive', this indicates that the rates are over graduated.

Question

Carry out the serial correlation test at lag 1 for Graduation A.

Graduation A

x	E_x^c	d_x	$\hat{\mu}_x$	$\mathring{\mu}_x$	$E_x^c \mathring{\mu}_x$	$z_x = \dfrac{d_x - E_x^c \mathring{\mu}_x}{\sqrt{E_x^c \mathring{\mu}_x}}$	z_x^2
30	70,000	39	0.000557	0.000388	27.16	2.27	5.16
31	66,672	43	0.000645	0.000429	28.60	2.69	7.25
32	68,375	34	0.000497	0.000474	32.41	0.28	0.08
33	65,420	31	0.000474	0.000524	34.28	−0.56	0.31
34	61,779	23	0.000372	0.000579	35.77	−2.14	4.56
35	66,091	50	0.000757	0.000640	42.30	1.18	1.40
36	68,514	48	0.000701	0.000708	48.51	−0.07	0.01
37	69,560	43	0.000618	0.000782	54.40	−1.55	2.39
38	65,000	48	0.000738	0.000865	56.23	−1.10	1.20
39	66,279	47	0.000709	0.000956	63.36	−2.06	4.23
40	67,300	62	0.000921	0.001056	71.07	−1.08	1.16
41	65,368	63	0.000964	0.001168	76.35	−1.53	2.33
42	65,391	84	0.001285	0.001291	84.42	−0.05	0.00
43	62,917	86	0.001367	0.001427	89.78	−0.40	0.16
44	66,537	120	0.001804	0.001577	104.93	1.47	2.16
45	62,302	121	0.001942	0.001743	108.59	1.19	1.42
46	62,145	122	0.001963	0.001926	119.69	0.21	0.04
47	63,856	162	0.002537	0.002129	135.95	2.23	4.99
48	61,097	151	0.002471	0.002353	143.76	0.60	0.36
49	61,110	184	0.003011	0.002601	158.95	1.99	3.95
Total		1,561			1,516.50		43.17

Solution

The mean of the individual standardised deviations is:

$$\bar{z} = \frac{1}{20}\sum_{X=30}^{49} Z_X = \frac{1}{20}(2.27 + 2.69 + \cdots + 1.99) = 0.18$$

The denominator of r_1 is:

$$\frac{1}{20}\sum_{X=30}^{49} ((Z_X - \bar{Z}))\text{^}2 = 2.13$$

The numerator of r_1 is:

$$\frac{1}{19}\sum_{X=30}^{48} (Z_X - \bar{Z})(Z_{X+1} - \bar{Z}) = 0.94$$

So:

$$r_1 = \frac{0.94}{2.13} = 0.44$$

And the value of the test statistic is

$$\sqrt{20}\ X\ 0.44 = 1.97$$

This is more than 1.6449, the upper 5% point of the standard normal distribution. So there is evidence of grouping of deviations of the same sign.

Question 3

A graduation of the mortality experience of the male population of a region of the United Kingdom has been carried out using a formula with 3 parameters. The following is an extract from the results.

Age nearest birthday	Actual number of deaths	Graduated mortality rate	Central exposed to risk	
x	θ_x	$\mathring{\mu}_x$	E_x^c	$E_x^c\mathring{\mu}_x$
14	3	0.00038	12,800	4.864
15	8	0.00043	15,300	6.579
16	5	0.00048	12,500	6.000
17	14	0.00053	15,000	7.950
18	17	0.00059	16,500	9.735
19	9	0.00066	10,100	6.666
20	15	0.00074	12,800	9.472
21	10	0.00083	13,700	11.371
22	10	0.00093	11,900	11.067
Total	91			73.704

Question 3

Continued…

i. Use the chi-squared test to test the adherence of the graduated rates to the data. State clearly the null hypothesis you are testing and comment on the result.
ii. Perform two other tests that detect different aspects of the adherence of the graduation to the data. For each test state clearly the features of the graduation that the test is able to detect, and comment on your results.

Solution

i. Chi-squared test

The null hypothesis is:

H_0 : the graduated rates are the true underlying mortality rates for the population

We calculate the individual standardized deviations at each age using the formula:

$$Z_x = \frac{\theta_x - E_x^c \dot{\mu}_x}{\sqrt{E_x^c \dot{\mu}_x}}$$

The ISDs are:

-0.845, 0.554, -0.408, 2.146, 2.328, 0.904, 1.796,-0.407, -0.321

The test statistic for the chi-squared test (based on unrounded z_x values) is:

$$\sum z_x^2 = 15.53$$

We now compare this with a χ^2 distribution. We were given data from 9 ages. Since the graduation was carried out using a formula with 3 parameters, we lose 3 degrees of freedom. So we are left with 6 degrees of freedom.

From the Tables, we see that the upper 5% point of χ_6^2 is 12.59. As the value of the test statistic is greater than this, we reject the null hypothesis and conclude that the graduated rates do not provide a good fit to the data. In particular, it looks like the graduated rates are too low for ages 17 to 20.

Solution

ii. Two other tests

You can take your pick here from the individual standardized deviations test, the signs test, the cumulative deviations test, the grouping of signs test and the serial correlation test.

The null hypothesis for all the tests is:
H_0 : the graduated rates are the true underlying mortality rates for the population

ISD Test
This is a good all round test that detects most of the problems that might be present in a graduation including any outliers.
For this test we compare the z_x values with a standard normal distribution:

	$(-\infty,-3)$	$(-3,-2)$	$(-2,-1)$	$(-1,0)$	$(0,1)$	$(1,2)$	$(2,3)$	$(3,\infty)$
Obs	0	0	0	4	2	1	2	0
Exp	0	0.18	1.26	3.06	3.06	1.26	0.18	0

Solution

Continued…

There are 4 things to consider here:

- Outliers-there are no ISDs greater than 3 in absolute value, which is good; however, with fewer than 20 age groups, we should be suspicious about any ISD greater than 2 in magnitude, and here we have 2 ISDs greater than 2.
- The balance of positive and negative deviations- this is OK
- Symmetry- the distribution is a bit positively skewed, which is not so good.
- Proportion of ISDs lying in the range $(-\frac{2}{3},\frac{2}{3})$ should be $\frac{1}{2}$ - it is $\frac{4}{9}$ here, which is OK.

- The graduated rates fail this test as the ISDs do not appear to be normally distributed. In particular, the graduated rates appear to be too low at ages 17 and 18.

Signs test

- This is a simple two-tailed test for overall bias.
- There should be roughly equal numbers of positive and negative ISDs. Under the null hypothesis, the number of positive deviations has a Binomial(9,0.5) distribution.
- We have 5 positives and 4 negatives, which is fine.
- So we do not reject the null hypothesis and we conclude that there is no evidence of overall bias in the graduated rates.

Solution

Continued…

Cumulative deviations test

- This is a two-tailed test for overall bias.
- The observed value of the test statistic is:

$$\frac{[\sum\theta_x - \sum(E_x^c\dot{\mu}_x)]}{\sqrt{\sum(E_x^c\dot{\mu}_x)}} = \frac{91 - 73.704}{\sqrt{73.704}} = 2.015$$

- For a test at the 5% significance level, we compare the value of the test statistic with the lower and upper 2.5% points of N(0,1) , ie with ± 1.96 . Since 2.015 is greater than 1.96, we reject the null hypothesis and conclude that the graduated rates are too low overall.

- Make sure that you don't do both of the signs test and the cumulative deviations test as they both test for the same thing.

Solution

Continued…

Grouping of signs test

- This is a one-tailed test that detects clumping of deviations of the same sign.
- The observed number of positive deviations is 5, and the observed number of negative deviations is 4.
- From the Tables, we find that the critical value is 1 and we reject the null hypothesis if the observed number of positive runs is less than or equal to this.
- The observed number of positive runs is 2, so we do not reject the null hypothesis in this case, and we conclude that there is no evidence of grouping of signs.

Solution

Continued…

Serial correlation test

- This is an alternative test for grouping of signs, but it takes much longer to carry out this test so it's recommended that you do don't do it unless you absolutely have to. Make sure that you don't carry out both the grouping of signs test and the lag-1 serial correlation test since they test for the same thing.
- This is a one-tailed test that detects clumping of deviations of the same sign.
- The serial correlation coefficient at lag 1 is:

$$r_1 = \frac{\frac{1}{8}\sum_{x=14}^{21}(z_x - \bar{z})(z_{x+1} - \bar{z})}{\frac{1}{9}\sum_{x=14}^{22}(z_x - \bar{z})^2} = \frac{0.2165}{1.3172} = 0.1643$$

- and the value of the test statistic is:

$$r_1\sqrt{m} = 0.1643X3 = 0.493$$

- As we are only testing for positive correlation, we compare the value of the test statistic with 1.6449, the upper 5% point of N(0,1) . We find that there is insufficient evidence to reject the null hypothesis or, in other words, there is no evidence of grouping of signs.

Comment

- The graduation has not fully taken into account the accident hump, ie the increase in mortality around the late teens and early twenties.

10 Methods for Graduation

- The process of graduation can be done using different methods.

- We will look at three methods of carrying out a graduation:
  - graduation by parametric formula
  - graduation by reference to a standard table
  - graduation using spline functions.

- The most appropriate method of graduation to use will depend on the quality of the data available and the purpose for which the graduated rates will be used.
- We will look at these methods in the next chapter

Recap

- Graduation is a technique used to get a smooth and justifiable mortality rates. (which is also called as Graduated Mortality Rate).
- The aims of graduation are: –
  - To produce a smooth set of rates that are suitable for a particular purpose.
  - To remove random sampling error.
  - To use information available from adjacent ages
- In mortality investigations and insurance studies, we often need to keep a check on the consistency of our data with other known experiences. It is important to see whether our recent experience is consistent with past experiences or with published life tables.
- By consistency, we cover two concepts: the shape of the mortality curve over the range of ages and the level of mortality rates.
- Three desirable features of a graduation are: – i. Smoothness; ii. Adherence to data & iii. Suitability for the purpose to hand.
- We need to make a balance between smoothing and adherence to data. At one extreme, we could easily smooth the crude estimates by ignoring the data altogether; we want to avoid such extremes since we want the graduation to be representative of the experience.
- If the graduation process results in rates that are smooth but show little adherence to the data, then we say that the rates are over-graduated.
- If insufficient smoothing has been carried out then it is called under-graduation.

Continued…

- Chi-squared test is a test for the overall fitness of fit. It can be used to assess whether the observed numbers of individuals in a specified categories are consistent with a model that predicts the expected numbers in each category.
- Standardized deviations test- The test looks at the distribution of the values of the standardized deviations.
- The signs test is a simple test for overall bias. In other words, this test checks whether the graduated rates are too high or too low.
- The cumulative deviations test detects overall bias or long runs of deviations of the same sign.
- Grouping of signs test is used mainly to test for over graduation.
- Serial correlations test tests for clumping of deviations of the same sign. If clumping is present, then the graduation has the wrong shape.